Finding the Beat in Music: Using Adaptive Oscillators

Kate M. Burgers

Jon Jacobsen, Advisor

Rachel Levy, Reader

May, 2011

Harvey Mudd College
Department of Mathematics

Abstract

The task of finding the beat in music is simple for most people, but surprisingly difficult to replicate in a robot. Progress in this problem has been made using various preprocessing techniques (Hitz, 2008; Tomic and Janata, 2008). However, a real-time method is not yet available. Methods using a class of oscillators called relay relaxation oscillators are promising. In particular, systems of forced Hopf oscillators (Large, 2000; Righetti et al., 2006) have been used with relative success. This work describes current methods of beat tracking and develops a new method that incorporates the best ideas from each existing method and removes the necessity for preprocessing.

Contents

Abstract **iii**

Acknowledgments **ix**

1 Introduction **1**
1.1 Musical Background 2

2 Mathematical Models **5**
2.1 Relay Relaxation Oscillator 5
2.2 Hopf Oscillator 7

3 Results **13**
3.1 Dependence on Parameters 14

4 Conclusion **17**
4.1 Future Work 18

Bibliography **19**

List of Figures

2.1 A typical relay relaxation oscillator 6
2.2 A phase portrait of the van der Pol oscillator 7
2.3 The behavior of a Hopf oscillator 8
2.4 The behavior of the Hopf oscillator with Large's adaptation 9
2.5 An illustration of a perturbation and basis vectors for an arbitrary limit cycle 10
2.6 The behavior of the Hopf oscillator with Righetti's adaptation 11
2.7 The behavior of the Hopf oscillator with the new adaptation 12

3.1 Square error of the output frequency vs. the coupling frequency 15
3.2 Square error of the output frequency vs the difference in internal and external frequency 16

Acknowledgments

I would like to thank Professors dePillis and Jacobsen for their guidance, and especially to Professor Jacobsen for his infinite patience as I navigated the thesis process. I would also like to thank Professor Levy for her help with MATLAB. My thanks also go to Ann Johnston for many late-night conversations, and to Jaron Kent-Dobias for converting my figures.

Chapter 1

Introduction

Most people are naturally able to recognize surprisingly complex information in a new piece of music nearly instantaneously. They can distinguish melodies from the basic chord structure and frequently begin tapping to the beat without a second thought. However, understanding how people are able to comprehend music with such speed and accuracy is not well understood and poses an interesting and challenging problem. My goal is not to understand the actual mechanism for such a deceptively simple task in people, but rather to create a new model that replicates the desired behavior.

The behavior I will focus on is *beat tracking*, which is perhaps the most second-nature behavior in human reactions to music. In fact, it is so instinctual, even infants will move to the beat of music, and there has been research showing that parrots are also capable of recognizing rhythm in music (Patel et al., 2008). My goal is to develop a mathematical model that is able to find the beat in music in real time.

Existing models use systems of oscillators with varying success and most of the work in beat tracking requires preprocessing, in which the model must "listen" to a large part of the music one or more times before coming up with an answer. For instance, Hitz (2008) uses adaptive oscillators to find tempo. Adaptive oscillators are directly affected by musical input, and will change their internal beat to catch up to the beat of the music. However, their process requires four preprocessing steps. Eck (2002) uses relaxation oscillators to detect the beat of simple rhythmic patterns. Relaxation oscillators will be discussed in more detail in Section 2.1. Eck's process was successful at beat prediction for a very specific rhythmic input. A different approach in Large (2000) uses Hopf oscillators, which are promising because they quickly return to the system's internal frequency, even after being perturbed. The Hopf oscillator will be discussed in more detail in Section 2.2. This approach seems to be fairly successful, although the given method requires a MIDI input with a small amount of preprocessing. I would like to adapt this method to work for any given sound input without preprocessing.

I propose using a family of dynamical systems called adaptive oscillators to find the beat in music. This oscillator would model the rhythm of a person walking across campus. Naturally, there is a pace which is most comfortable for each person, which is their natural frequency. However, when they walk past a dorm playing some upbeat music, their natural tendency is to walk faster so their steps match the beat of the music. Thus, the oscillator will have the properties that it has a natural internal frequency, which it will revert to when there is no external forcing, and it will be influenced by an arbitrary input. In addition, it may be possible that the oscillator's natural frequency can be changed by prolonged exposure to an input with a different frequency.

1.1 Musical Background

We can better understand the meaning of rhythm and beat from a music theoretic perspective. Western culture has developed musical notation which precisely defines dynamics, tempo and rhythm. Dynamics refer to how loud or quiet a piece of music is and is shown by the amplitude of a digital input. Most music has a wide range of dynamics, but this does not seem to affect a person's ability to find a beat accurately. In contrast, changing dynamic levels can create confusion for a robot which adapts based on the amplitude of the input forcing.

Half of the beat-tracking problem is matching the period or tempo of the music. The tempo for a piece of music is generally set at the beginning of a piece of music between 40 beats per minute and 200 beats per minute. However, because people do not play music at precisely the tempo specified, and may not keep a steady beat, the given tempo is merely a suggestion and is not enough information for a computer to match the tempo of a given piece of music. Furthermore, a composer may choose to change the tempo part way through the music, either gradually or abruptly. Therefore, the beat found at the beginning of a piece may not be the same beat for the whole piece, and a beat-tracking model should be able to adapt to changes in the music.

The second half of beat tracking is matching the phase or the downbeat. A time signature at the beginning of the music consists of two numbers. The top number lets a performer know how many beats are in the measure and the bottom number tells the performer what kind of note is considered the main beat. For example, a common time signature, $\frac{3}{4}$, indicates that the quarter note gets the beat, and there are three beats per measure. The downbeat is defined as the first beat of a measure, so we want our model to reach its maximum in each measure at the onset of the downbeat. This goal is complicated by the fact that there does not necessarily have to be a note sounding on the downbeat, and there can be notes in between main beats. Thus a beat-tracking model should be able to distinguish between a tempo change and notes that occur on the off-beat.

Chapter 2

Mathematical Models

In order to match human behavior, we want our model to have some basic properties.

1. The model should have an "internal beat" that it tends to when there is no music present.

2. After prolonged exposure, the model's internal beat should adapt to match the rhythm of the music ("find the beat").

3. The model should be able to adapt to music in real-time.

This chapter will discuss several models which fulfill one or more of these properties.

2.1 Relay Relaxation Oscillator

Relaxation oscillators are a class of oscillators, which operate on two separate timescales, sometimes called a bistable subsystem, which consists of a fast timescale and a negative integral action, in which the system slowly builds in one direction until it suddenly switches states. An example of a relaxation oscillator is a seesaw with a weight on one end and a cup with water slowly dripping into it on the other end. As the water drips into the cup, it slowly tips the seesaw until a certain point when the cup suddenly dumps all of its water and the seesaw resets. In this example, the seesaw corresponds to the bistable subsystem, the seesaw slowly tipping as water drips into the cup corresponds to the negative integral action and the action of the water dumping corresponds to the fast time scale. Furthermore,

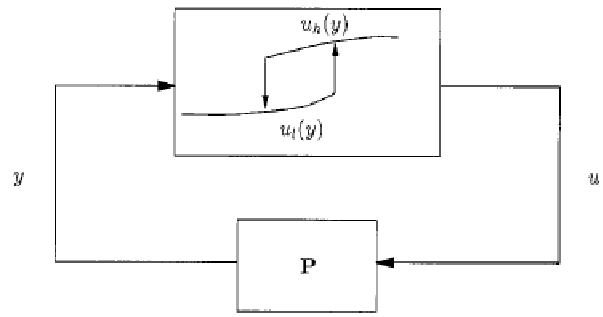


Figure 2.1 This is a model of a typical relay relaxation oscillator. It is characterized by the interaction of a system exhibiting hysteresis and a linear system P. Originally published in Varigonda and Georgiou (2001).

relaxation oscillators have stable limit cycles, or self-sustaining oscillations that are robust to perturbations.

A relay relaxation oscillator is a relaxation oscillator that exhibits relay hysteresis (Varigonda and Georgiou, 2001). Hysteresis is characterized by the simultaneous existence of two or more stable states and the possibility for the system to jump from one stable state to another as the system evolves in time or space. Relay hysteresis means that there is a delay in the switching for the hysteresis. The setup of a typical relay relaxation oscillator is shown in Figure 2.1.

2.1.1 Van der Pol Oscillator

The most common example of a relaxation oscillator is the van der Pol oscillator, which is described by

$$\epsilon u' = y - \frac{u^3}{3} + u \tag{2.1}$$

$$y' = -u. \tag{2.2}$$

By setting $y' = 0$ we find that the y-nullcline is $u = 0$, and by setting $u' = 0$ we find that the u-nullcline is $y = \frac{u^3}{3} - u$. If ϵ is small, then $u' >> y'$ unless y is close to the u-nullcline. In this example, the slow time scale occurs near the u-nullcline, as shown in Figure 2.2.

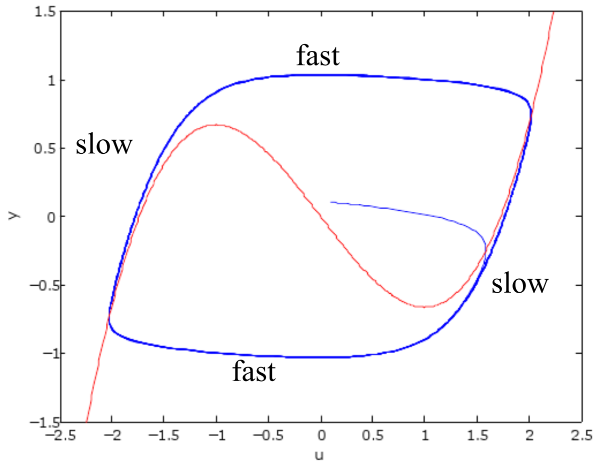


Figure 2.2 This is a phase portrait for the van der Pol oscillator for small ϵ (blue). The van der Pol oscillator operates on a slow timescale when it is near the u-nullcline, $y = \frac{u^3}{3} - u$ (red).

2.2 Hopf Oscillator

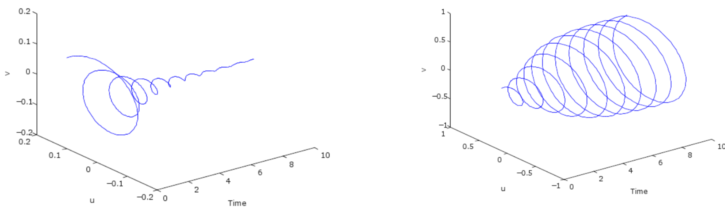
A model that I will be considering in detail is the Hopf oscillator. The basic, unforced Hopf oscillator is given by

$$u' = \alpha u - \omega v - u^3 - uv^2 \tag{2.3}$$
$$v' = \alpha v + \omega u - u^2 v - v^3, \tag{2.4}$$

where α is the strength of the oscillator and ω is the natural frequency of the oscillator. For this model, we can show that the amplitude and phase are independent of one another. Furthermore, there is an equilibrium point at $(0, 0)$, which is stable when $\alpha < 0$, and unstable when $\alpha > 0$. In fact, the Jacobian at the equilibrium point is

$$\begin{bmatrix} \alpha & -\omega \\ \omega & \alpha \end{bmatrix}$$

with eigenvalue $\lambda = \alpha \pm i\omega$. Thus by the Hartman-Grobman theorem, the equilibrium point $(0, 0)$ is locally asymptotically stable for $\alpha < 0$ and unstable for $\alpha > 0$. When $\alpha > 0$ the system has a stable limit cycle with amplitude $\sqrt{\alpha}$. That is, there is a Hopf bifurcation at $\alpha = 0$. Two typical examples are shown in Figure 2.3.



a. $\alpha < 0$ **b.** $\alpha > 0$

Figure 2.3 The behavior of a Hopf oscillator with varying values for α.

It will be useful later in this section to refer to the polar form of the Hopf oscillator. We can derive the polar form using the identities $x = r\cos\phi$ and $y = r\sin\phi$. Plugging these into the previous equation and solving for r' and ϕ' gives

$$r' = r(\alpha - r^2) \tag{2.5}$$
$$\phi' = \omega. \tag{2.6}$$

It is clear from the polar form that the amplitude and phase are independent, and that there is a limit cycle when $\alpha > 0$ with radius $r = \sqrt{\alpha}$. Note that for $\alpha < 0$, $r' < 0$ for all t, so all solutions are driven to the stable equilibrium point $(0,0)$. That is, $(0,0)$ is globally asymptotically stable.

In order for any oscillator to find a beat in music, it needs to have a way of "hearing" the music. For our system to "hear" music, we need to add a forcing term. In addition to this forcing term, Large (2000) proposes that many oscillators in competition will be more likely to accurately find the beat than a single oscillator. Therefore, he proposed a forced Hopf oscillator system with N oscillators affecting one another. The system is defined by

$$u_n' = \eta_n \mathrm{Re}\,[F(t)]e^{\frac{\kappa_n(u_n - \sqrt{u_n^2+v_n^2})}{\sqrt{u_n^2+v_n^2}}} - \omega_n v_n - u_n^3 - u_n v_n^2 + \sum_{m\neq n} \gamma_{m,n}(u_n u_m^2 + u_n v_m^2) \tag{2.7a}$$

$$v_n' = \eta_n \mathrm{Im}\,[F(t)]e^{\frac{\kappa_n(u_n - \sqrt{u_n^2+v_n^2})}{\sqrt{u_n^2+v_n^2}}} + \omega_n u_n - v_n^3 - u_n^2 v_n + \sum_{m\neq n} \gamma_{m,n}(v_n u_m^2 + v_n v_m^2), \tag{2.7b}$$

where η is the coupling strength, or the amount of influence the external

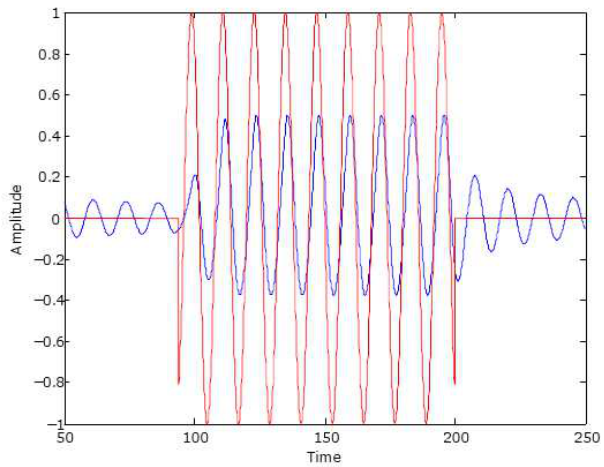


Figure 2.4 The behavior of Large's adaptation (blue) with input $F(t) = \sin \frac{\pi t}{6}$ (red) starting at $t = 95$ and ending at time $t = 200$.

forcing exerts; $F(t)$ is the forcing term; κ is dependent on ω, the frequency of the oscillator; γ is a matrix of inhibition parameters whose entries tell how much each oscillator influences the others; and $n \in \{1, 2, \ldots, N\}$. Note that this is a system of $2N$ equations.

It will also be useful to consider this model in polar form. The polar form is given by

$$r'_n = r_n \left(\eta_n F(t) e^{\frac{\kappa_n(\Re(\phi_n) - |\phi_n|)}{|\phi_n|}} - r_n^2 \right) - \sum_{m \neq n} \gamma_{m,n} r_n r_m^2 \tag{2.8}$$

$$\phi'_n = \omega_n - \eta_n F(t) e^{\frac{\kappa_n(\Re(\phi_n) - |\phi_n|)}{|\phi_n|}} \sin \phi_n. \tag{2.9}$$

It is now clear that adding a forcing term to the Hopf oscillator system creates a dependence between the amplitude and phase of the system. A sample solution for $N = 1$, forcing input $F(t) = sin \frac{\pi t}{6}$ beginning at $t = 95$ and internal frequency $\omega = 0.5$ is shown in Figure 2.4. Notice that the model is periodic when there is no forcing (at the internal beat), adapts to match the frequency of the input when it starts (finds the beat) and finally goes back to its original frequency after the forcing disappears.

Righetti, Buchli, and Ijspeert (2006) give a different twist to the original model: they add a third equation that allows the natural frequency of the oscillator to be a variable. This variation allows the natural frequency of

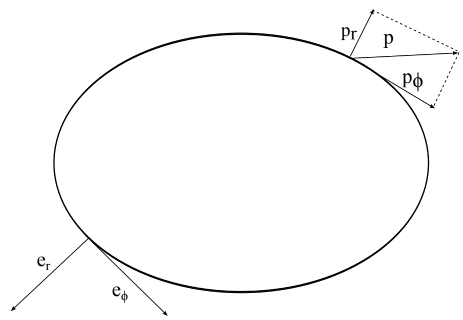


Figure 2.5 An illustration of a perturbation and basis vectors for an arbitrary limit cycle. The dominant effect on the limit cycle is $p_\phi = p \cdot e_\phi$. Adapted from Righetti et al. (2006).

the oscillator to explicitly change after prolonged exposure to an input. The third equation will be derived from the forced Hopf system given by

$$r' = (\mu - r^2)r + \pm F \cos\phi \tag{2.10a}$$

$$\phi' = \omega \pm \frac{\epsilon}{r} F \sin\phi. \tag{2.10b}$$

The adaptation rule for the frequency is derived by analyzing the perturbation of the external input on the oscillator, so it will have the general form

$$\omega' = f(\omega, r, \phi, F).$$

If we consider the perturbation as a vector $\vec{P}$ in phase space, then we can write its components with respect to a radial base vector $\vec{e}_r$ and a tangential base vector $\vec{e}_\phi$, as shown in Figure 2.5. Recall that the Hopf oscillator has a stable limit cycle, so any perturbations will not affect its radius in the long term. Therefore, the external input will only affect the phase of the oscillator. The influence of the phase will be given by the projection of $\vec{P}$ onto $\vec{e}_\phi$:

$$p_\phi = \vec{P} \cdot \vec{e}_\phi.$$

For the Hopf oscillator with the forcing given in Equation 2.10 we have

$$p_\phi = \frac{\pm\epsilon}{r} F \sin\phi.$$

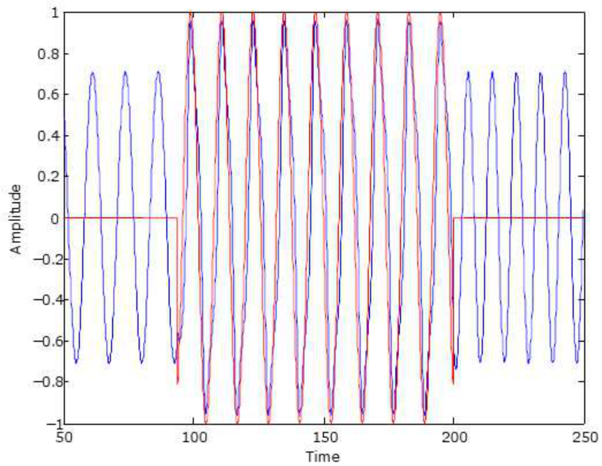


Figure 2.6 The behavior of Righetti's adaptation (blue) with input $F(t) = \sin \frac{\pi t}{6}$ (red) starting at $t = 95$ and ending at time $t = 200$.

Thus for this Hopf oscillator system, the impact of a forcing F on the natural frequency ω is

$$\omega' = \pm \epsilon F \frac{v}{\sqrt{u^2 + v^2}}$$

for the frequency, where the sign is determined by the natural direction of the oscillation and ϵ is the strength of the forcing. Therefore, the forced version of the Hopf oscillator should look like

$$u' = \alpha u - \omega v - u^3 - uv^2 + \epsilon F(t) \tag{2.11}$$

$$v' = \alpha v + \omega u - u^2 v - v^3 \tag{2.12}$$

$$\omega' = -\epsilon F(t) \frac{v}{\sqrt{u^2 + v^2}}. \tag{2.13}$$

This model is more effective than Large's model for beat tracking, but does not allow for multiple oscillators to be running at the same time. A sample solution for Righetti's model is shown in Figure 2.6

We can combine Large's multiple oscillator model with Righetti's adaptive frequency model to create a new model that has all the benefits of both previous models. Using Righetti's derivation for ω' and the system given in Equation 2.7, we find that for the n^{th} oscillator,

$$\omega_n' = -\eta_n F(t) e^{\frac{\kappa_n \mathrm{Re}\,(\phi_n)}{|\phi_n|} - \kappa_n} v_n.$$

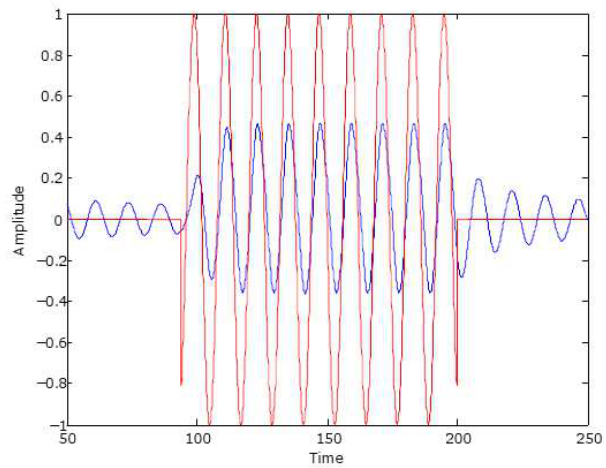


Figure 2.7 The behavior of the new adaptation (blue) with input $F(t) = \sin \frac{\pi t}{6}$ (red) starting at $t = 95$ and ending at time $t = 200$.

A sample solution for one oscillator and input $F(t) = \sin \frac{\pi t}{6}$ starting at $t = 95$ and ending at time $t = 200$ is given in Figure 2.7.

Chapter 3

Results

To test how accurately each model finds the beat, we compute numerical solutions to the systems of differential equation using MATLAB. For each model, simulations were run for sinusoidal inputs $F(t) = \sin \omega_{in} t$ with angular frequencies ranging from $0 \leq \omega_{in} \leq 5$. Furthermore, for each input, the parameters ω, α, ϵ, and η were varied. To make comparison across models easier, we let $\alpha = \epsilon = \eta$ for each trial. The range of values for the parameters are $0 \leq \omega \leq 5$ and $0 \leq \alpha \leq 3.5$. For both Large's model and the new model, we use one oscillator ($N = 1$).

Once we have a solution to a system, we can use a fast Fourier transform to extract the frequency of the output function, f_{out}, and compare it to the input frequency $f_{in} = \frac{\omega_{in}}{2\pi}$. It is important to note that the fast Fourier transform breaks a function into its frequency components, and it is most likely that the output function will have more than one frequency component. Therefore, we will define the output frequency as the nonnegative frequency with the maximum amplitude. We define the error in frequency as the Euclidean distance:

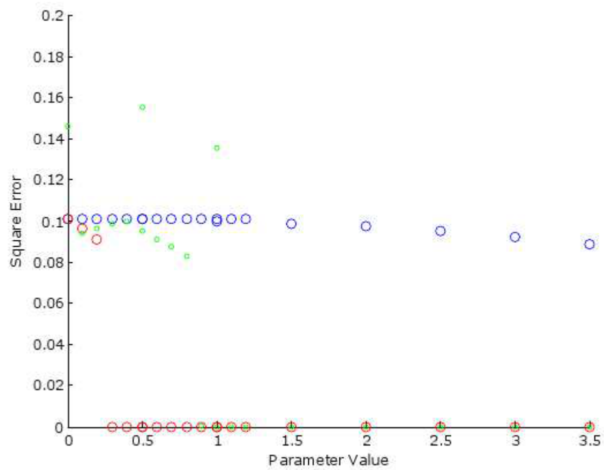
$$E = |f_{in} - f_{out}|^2.$$

We find that for the given range of parameters, Large's model has the largest average error, and Righetti's model has the smallest average error. That means when $N = 1$, the new model seems to be an improvement on Large's model, but does not seem improve Righetti's model.

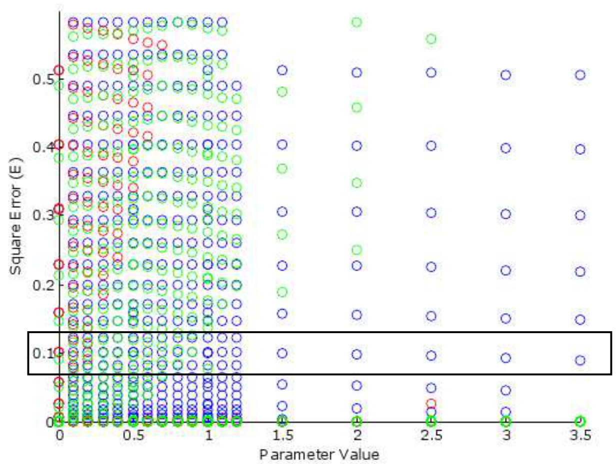
3.1 Dependence on Parameters

Given the number of parameters in our models, we would like to know how much the error depends on each parameter. The larger the coupling strength, the more the external forcing is influencing the oscillator. Therefore, we expected that coupling strength (η in Large's model and ϵ in Righetti's model) would be negatively correlated with error. However, as shown in Figure 3.1, there appears to be no correlation between the coupling strength and error. This means that as long as we choose a positive coupling strength, we should not expect the output to change drastically.

We expect that if we begin with an internal frequency ω that is close to the input's angular frequency ω_{in}, then it will take less time for the oscillator to adapt to the input. If the oscillator is adapting quickly, we should see less error in the output frequency f_{out}. We can see in Figure 3.2 that this appear to be the case. Therefore, it is important for us to start with an internal frequency which is close to the input's angular frequency. In this regard, it may be better to use the new adaptation with many oscillators each with a different internal frequency than Righetti's adaptation.



a. The square error of the output frequency vs. the coupling strength for $\omega_{in} - \omega_{out} = 2$.



b. The square error of the output frequency vs. the coupling strength for all trials. Notice the data for Large's model and the new model in Figure 3.1a falls almost entirely within the box, indicating that there is no correlation.

Figure 3.1 The square error of the output frequency for Large's adaptation (blue), Righetti's adaptation (red) and the new adaptation (green) as a function of the coupling frequency, η for Large's and the new adaptation and ϵ for Righetti's adaptation. Notice that there seems to be no correlation between the coupling frequency and the error E.

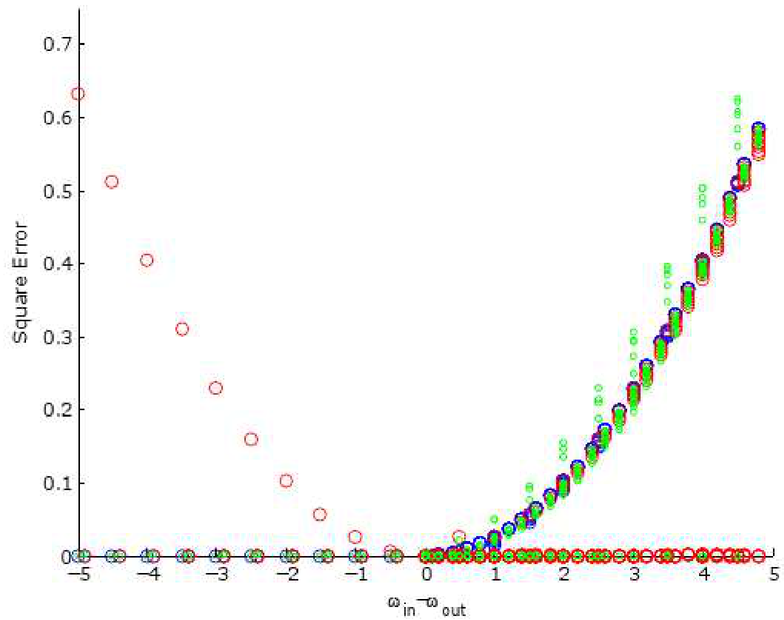


Figure 3.2 The square error of the output frequency for Large's adaptation (blue), Righetti's adaptation (red) and the new adaptation (green) as a function of the difference in internal frequency and external frequency. Notice that for Large's adaptation and the new adaptation, as $\omega_{in} - \omega_{out}$ increases, the error E also increases. However, for Righetti's model, the opposite correlation seems to hold.

Chapter 4

Conclusion

Our goal was to find a model which could adapt to find the beat in music in real time. We added the additional requirements that our model should have an internal beat which could adapt after prolonged exposure to music. We proposed that adaptive oscillators would be best suited for the task. More specifically, we considered adaptations to the Hopf oscillator, which is particularly promising because it has stable limit cycles. Large proposed that we couple N competing Hopf oscillators together so that the oscillator which does the best job of finding the beat will continue to oscillate while the others tend to zero. Righetti instead proposed that we add a third equation to the Hopf system which allows the frequency ω to vary in time. Finally, we combine both of these ideas into a new model:

$$u_n' = \eta_n \mathrm{Re}\,[F(t)] e^{\frac{\kappa_n (u_n - \sqrt{u_n^2+v_n^2})}{\sqrt{u_n^2+v_n^2}}} - \omega_n v_n - u_n^3 - u_n v_n^2 + \sum_{m\neq n} \gamma_{m,n}(u_n u_m^2 + u_n v_m^2)$$

$$v_n' = \eta_n \mathrm{Im}\,[F(t)] e^{\frac{\kappa_n (u_n - \sqrt{u_n^2+v_n^2})}{\sqrt{u_n^2+v_n^2}}} + \omega_n u_n - v_n^3 - u_n^2 v_n + \sum_{m\neq n} \gamma_{m,n}(v_n u_m^2 + v_n v_m^2)$$

$$\omega_n' = -\eta_n F(t) e^{\frac{\kappa_n \mathrm{Re}\,(\phi_n)}{|\phi_n|} - \kappa_n} v_n,$$

where ϕ_n is the phase of the complex number $u_n + iv_n$, $F(t)$ is the forcing, and $n \in \{1, 2, \ldots, N\}$. Notice that this is a system with $3N$ equations.

We found that if we let $N = 1$, then Righetti's model has the least error when matching the frequency of the input. For all three models, the error is directly proportional to the difference between the internal frequency and the input frequency, so it seems likely that if we let $N > 1$, with each oscillator at a different internal frequency, the new model may perform better.

On the other hand, there seems to be no correlation between the coupling strength and the error.

4.1 Future Work

It would be interesting to consider how Large's adaptation and the new adaptation work when there are more than one oscillator, especially when every oscillator has a different internal frequency. A challenge in interpreting the results for $N > 1$ is deciding whether to use data from multiple oscillators or one oscillator. Furthermore, one must decide autonomously which oscillator(s) are giving accurate information, and which oscillators should be ignored.

In this paper, we restricted our attention to simple sine wave inputs. In the future, it would be interesting to consider more complicated inputs such as sums of sine waves, or real music.

Finally, we have focused on each model's ability to find the tempo of an input. However, in beat tracking, matching the phase of the music is just as important as knowing how fast it is moving. Therefore, considering each oscillator's ability to phase lock, and further adapting the given models to improve in phase locking would be very useful.

Bibliography

Bregman, Micah R. 2003. Adaptive oscillator models of metrical perception. Pomona College Senior Thesis. Senior Thesis.

Eck, Douglas. 2002. Finding downbeats with a relaxation oscillator. *Psychological Research* 66:18–25.

Hitz, Aïsha. 2008. *Synchronization of Movements of a Real Humanoid Robot with Music*. Master's thesis, École Polytechnique Fédérale de Lausanne. Master Project.

Large, Edward W. 2000. On synchronizing movements to music. *Human Movement Science* 19:527–566.

Large, Edward W., and John F. Kolen. 1994. Resonance and the perception of musical meter. *Connection Science* 6(2–3):177–208.

Large, Edward W., and Caroline Palmer. 2002. Perceiving temporal regularity in music. *Cognitive Science* 26:1–37.

Lots, Inbal Sapira, and Lewi Stone. 2008. Perception of musical consonance and dissonance: An outcome of neural synchronization. *Journal of the Royal Society Interface* 5:1429–1434.

Patel, Aniruddh D., John R. Iversen, Micah R. Bregman, Irena Schulz, and Charles Schulz. 2008. Investigating the human-specificity of synchronization to music. *Proceedings of the 10th International Conference on Music Perception and Cognition* 1–5.

Righetti, Ludovic, Jonas Buchli, and Auke Jan Ijspeert. 2006. Dynamic Hebbian learning in adaptive frequency oscillators. *Physica D* 216:269–281.

Strogatz, Steven H., and Ian Stewart. 1993. Coupled oscillators and biological synchronization: A subtle mathematical thread connects clocks, ambling elephants, brain rhythms and the onset of chaos. *Scientific American* 102–109.

Toiviainen, Petri, and Joel S. Snyder. 2003. Tapping to Bach: Resonance-based modeling of pulse. *Music Perception* 21(1):43–80.

Tomic, Stefan T., and Petr Janata. 2008. Beyond the beat: Modeling metric structure in music and performance. *The Journal of the Acoustical Society of America* 124(6):4024–4041.

Varigonda, Subbarao, and Tryphon T. Georgiou. 2001. Dynamics of relay relaxation oscillators. *IEEE Transactions on Automatic Control* 46(1):65–77.

Yamanishi, Jun-ichi, Mitsuo Kawato, and Ryoji Suzuki. 1980. Two coupled oscillators as a model for the coordinated finger tapping by both hands. *Biological Cybermetrics* 37:219–225.